ED 316 557 TM 014 507

AUTHOR Sawilowsky, Shlomo S.; Markman, Barry S.
TITLE GUIL"t".
PUB DATE 89
NOTE 7p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Evaluation Problems; Higher Education; Mathematical Models; Monte Carlo Methods; *Sample Size
IDENTIFIERS Power (Statistics); Robustness; *T Test; Type I Errors

ABSTRACT

A problem that often surfaces in the use of the "t"-test is the absence of critical values for common sample sizes. This problem may cause "guilt" on the part of the professor who must advise students when they encounter discrepancies between their own calculations of the degree of freedom and critical values provided in popular statistics and research textbooks. To avoid hazards of interpolating, professors often suggest using a higher or lower critical value, which usually differs only in the thousandths place. The current study is a Monte Carlo investigation comparing the robustness and power properties of the independent means "t"-test when using the correct critical value versus the approximate values under normal and various non-normal distributions at different alpha levels. The results indicate that the lower critical value allows increased chances of making a Type I error. In the worst case, use of the higher critical value resulted in a modest (1.25%) loss in power. Two data tables are provided. (TJH)

GUILt

SHLOMO S. SAWILOWSKY

Educational Evaluation and Research
College of Education
Wayne State University

BARRY S. MARKMAN

Educational Psychology
College of Education
Wayne State University

ABSTRACT

A problem that often surfaces in the use of the t-test is the absence of critical values for common sample sizes. To avoid the hazards in interpolating, professors often suggest using the higher or lower critical value, which usually differs only in the thousandths place. The current study is a Monte Carlo investigation comparing the robustness and power properties of the independent means t-test when using the correct critical value vs. the approximate values under normal and various nonnormal distributions at different alpha levels. The results indicated that the lower critical value allows increased chances of making a Type 1 error. In the worst case, use of the higher critical value resulted in a modest (6.25%) loss in power.

G U I L $\underline{t}$

1. Introduction

According to English and English (1965), guilt is defined as a violation of principles that leads to a regretful feeling of lessened personal worth. Guilt, in contradistinction to shame, relies on both internal and external sanctions. For the purposes of the current stuoy the focus will be on internal sanctions.

As faculty and statistical advisors, we implore our students to avoid errors in computations by not rounding until the final step has been completed. For example, after applying the conceptual formula for population variance

$$\sigma^2 = \frac{\Sigma(X-\mu)^2}{N} \qquad (1)$$

on a data set of N = 5, students may obtain different results at the 1/10ths place due to rounding errors. This can be attributed to the five rounding opportunities that occur in subtracting $X-\mu$. To avoid this the computational formula is introduced:

$$\sigma^2 = \frac{\Sigma x^2 - \frac{(\Sigma X)^2}{N}}{N} \qquad (2)$$

and the students are instructed not to round until the last step. Similarly, in calculating the Pearson correlation coefficient, students are informed thai esults should be significant to the 1/100ths or even the 1/1000ths place.

1.1 The Problem

We teach this and then we try to sleep at night. Why should sleep escape us? Guilt! We may toss and turn because after the quest for precision in the descriptive statistics unit the pedagogy changes in the inferential statistics unit.

Consider the $\underline{t}$-test. Suppose students are given a data set that requires calculating a two independent means $\underline{t}$-test on response data for Condition A and Condition B in a balanced layout with N = 19 per group. The degrees of freedom, N1+N2-2, is 36. A survey of over one hundred popular statistics and research textbooks shows tabled values of $\underline{t}$ for one through 30, 40, 60, 120, and ∞ degrees of freedom. (A few textbooks also included 45 d.f.)

The students invariably ask which tabled critical value in their textbook should be used since 36 d.f. is not there. Of course, the best answer is to advise the students to obtain the correct critical value from tables available in the library, which apparently is an extremely taxing proposition for many students. Some faculty suggest interpolating between the tabled values for 30 and 40. This process is somewhat less than satisfying for two reasons. First, since the distribution is not uniform the interpolation does not follow the law of isomorphism. Related to this is the fact that the descent of the curve is not as rapid for α = .010 as it is when α = .050. Second, it is somewhat unnerving to have students using different critical values based on their own methods of interpolation, which could lead to different statistical conclusions.

Another less than satisfactory approach would be to avoid the $\underline{t}$ tables and compare the obtained $\underline{t}$ value to the critical Z, since at larger sample sizes $\underline{t}$ approaches the value of Z. All textbooks reviewed contained exter ive Z tables. Hays (1988) stated:

> If the population is truly normal, even 40 or so cases permit a fairly accurate use of the normal tables in confidence intervals or tests for a mean. If really good accuracy is desired in determining interval probabilities, the *t* distribution should be used even when the sample size is around 100 cases. Beyond this number of cases, the normal probabilities are extremely close to the exact *t* probabilities. (p. 292)

Using this method would avoid the problems associated with interpolating, but the explanation for using Z values in a *t*-test could never be satisfactory.

1.2 The Solution?

There are two other, more prevalent, solutions. The critical *t* value for 30 degrees of freedom or for 40 degrees of freedom, which are given in all textbooks reviewed, could be used. The *t* statistic associated with 30 degrees of freedom will be conservative, while the *t* statistic associated with 40 degrees of freedom will be liberal. An inspection of the *t* tables indicates only minor differences at sample sizes of 30 and 40, differences mainly in the 1/100ths place. How does the imprecision of using the wrong critical value affect the robustness of Type I error and power properties of the *t*-test?

2. Methodology

The current study is a Monte Carlo investigation of the properties of the two sample independent means *t*-test under the normal and various nonnormal distributions when the wrong critical values have been used. A Fortran program was written in MicroSoft Fortran 5.0 for an Intel 80386 IBM compatible computer with an Intel 80387 numeric coprocessor, accessing the PC version of IMSL (1987) RNUN, RNCHI, RNCHY, RNEXP, and RNNOA subroutines.

To examine the robustness characteristics, the independent means *t*-test was applied to two groups of random variates drawn from normal, uniform, chi-square (d.f.=1), cauchy, and exponential distributions. In addition, random variates were drawn from a multimodal lumpy distribution described by Micceri (1989). These distributions were selected since they represent a wide variety in terms of skew, tail weights, and asymmetry. The distribution described by Micceri is a real world data distribution of particular interest due to its extreme nonnormal shape. Random variates were drawn from a population and the *t*-test using the correct and approximate critical values was applied at both the .050 and .010 alpha level. A fail to reject vs. reject decision was recorded. This was repeated for 10,000 repetitions for each condition studied.

In the second part of this study the power pr[illegible]ies of the *t* statistic when using the correct and approximate critical values was studied. A treatment effect of shift in location parameter was modeled by adding a constant, c, to each of the observations of one group. The c value for each distribution was selected to obtain a power level of .500 for the correct *t* value when nominal alpha was .100. That same constant was used to obtain the power at the .050 and .010 alpha levels as well.

3. Robustness Results

The proportion of rejections for each distribution is depicted below in Table 1. Entries are provided for the higher *t* (t_H), *t*, and the lower *t* (t_L) at the .050 and .010 alpha level for each of the distributions investigated.

(Place Table 1 about here)

The only distribution in which the incorrect critical *t* values gave divergent results was the exponential, and even with this distribution it was only at the .050 alpha level that non-acceptable results were obtained. Using the higher *t* value generated a relatively conservative test, with actual alpha at .043. The lower *t* value rejected more than the nominal .050, with the actual rate of rejections at .055.

4. Power Results

The power results are presented below, in Table 2. For each distribution, the rejections at various alpha levels are given for t_H, t, and the t_L.

(Place Table 2 about here)

For the various alpha levels investigated, the power results in Table 2 indicate that t_L does not gain any appreciable power over the correct t under any of the distributions investigated. Therefore, at the cost of being slightly liberal, the use of the lower t value fails to return any power advantage. However, the power results for t_H demonstrated an appreciable loss of power, particularly under the normal (6.25% loss) and exponential distributions (6% loss) at the .010 alpha level. However, at the .050 alpha levels the power loss is less severe.

5. Discussion and Recommendation

The use of t_L (t value for the smaller sample size n=30) is not the best choice. The marginal increase in power is moot; the slightly increased Type I error rate makes it slightly dangerous. The use of t_H (t value for the larger sample size n=40) is a better choice. It is conservative (rejecting less than nominal alpha), and except under the smallest alpha level under two of the six distributions studied, the cost in power was negligible.

Nevertheless, it surely is a basic principle of education that students transfer learned behaviors beyond the classroom. The temptation to use t_H, with its slight loss of power, should be weighed against the abundant literature in education and psychology that suggests treatment effects in these disciplines are often quite small. In such situations the loss of any power to detect a false null hypothesis is exacerbated.

As a solution to the problem, we recommend that textbook authors include the missing critical values, at least for N = 31 through 39 and 41 through 59. Until such time, however, we suggest that the *professor go to the library*, gather the missing critical values, and include them as curricular materials given to the students. At least that way, we should be able to get a reasonable night's rest.

6. References

English, H. B., and English, A. C. (1965). A comprehensive dictionary of psychological and psychoanalytical terms. N.Y.: David McKay Company, Inc.

Hays, W. L. (1988). Statistics. (4th ed.). N.Y.: Holt, Rinehart, and Winston, Inc.

IMSL (1987). IMSL STAT/Library User's Manual: Fortran subroutines for statistical analysis. Version 1.0. Houston: Author.

Micceri, T. (1989). The unicorn, the normal curve, and other improbably creatures. Psychological Bulletin, 105, No. 1., 156-166.

TABLE 1.
Type I Error Rates When Using The Higher t (t_H), t, And The Lower t (t_L) Critical Values For Normal And Nonnormal Distributions, With Alpha=.050 And .010, And N=19 In Each Group.

Critical Value	Normal	ML	Uniform	Expn	ChiSq	Cauchy
			Distribution			
t_H	.050	.049	.049	.043	.039	.018
	.011	.009	.010	.008	.005	.002
t	.050	.050	.049	.051	.041	.019
	.011	.010	.010	.009	.005	.002
t_L	.051	.052	.050	.055	.042	.019
	.011	.012	.011	.009	.005	.002

Note: Expn = Exponential, ChiSq = Chi Square (d.f.=1), ML = Multimodal Lumpy.

TABLE 2.
Power Rates When Using The Higher t (t_H), t, And The Lower t (t_L) Critical Values For Normal And Nonnormal Distributions, With Alpha=.100, .050, And .010, And N=19 In Each Group.

Crit. Value		Normal	ML	Uniform	Expn	ChiSq	Cauchy
				Distribution			
t_H	.100	.494	.495	.494	.496	.497	.494
	.050	.368	.366	.348	.373	.388	.411
	.010	.160	.158	.139	.167	.194	.286
t	.100	.500	.500	.500	.500	.500	.500
	.050	.374	.368	.358	.380	.396	.416
	.010	.172	.160	.145	.177	.198	.295
t_L	.100	.501	.501	.503	.502	.501	.500
	.050	.376	.369	.362	.387	.402	.419
	.010	.172	.161	.148	.181	.199	.296

Note: Expn = Exponential, ChiSq = Chi Square (d.f.=1), ML = Multimodal Lumpy, Crit. Value = Critical Value.